

# Evidence for chromatic edge detectors in human vision using classification images

William McIlhagga

Bradford School of Optometry & Vision Science,  
University of Bradford, Bradford, UK



Kathy T. Mullen

McGill Vision Research, Department of Ophthalmology,  
Montreal, Quebec, Canada



Edge detection plays an important role in human vision, and although it is clear that there are luminance edge detectors, it is not known whether there are chromatic edge detectors as well. We showed observers a horizontal edge blurred by a Gaussian filter (with widths of  $\sigma = 0.1125, 0.225$ , or  $0.45^\circ$ ) embedded in blurred Brown noise. Observers had to choose which of two stimuli contained the edge. Brown noise was used in preference to white noise to reveal localized edge detectors. Edges and noise were defined by either luminance or chromatic contrast (isoluminant L/M and S-cone opponent). Classification image analysis was applied to observer responses. In this analysis, the random components of the stimulus are correlated with observer responses to reveal a template that shows how observers weighted different parts of the stimulus to arrive at their decision. We found classification images for both luminance and isoluminant chromatic stimuli that had shapes very similar to derivatives of Gaussian filters. The widths of these classification images tracked the widths of the edges, but the chromatic edge classification images were wider than the luminance ones. These results are consistent with edge detection filters sensitive to luminance contrast and isoluminant chromatic contrast.

## Introduction

When the world is projected onto the retina, physical object boundaries produce changes in the intensity of the retinal image; that is, edges. Some changes in intensity are due to object boundaries, but other physical processes such as surface orientation and scene illumination also create changes in image intensity (Barrow & Tenenbaum, 1981; Marr, 1982). Although changes in luminance intensity can be created by any of the aforementioned three physical processes, changes in color predominantly arise from object boundaries, due to changes in surface reflectance. Thus, color changes

may be used as indicators of object boundaries, whereas their absence is an indication of surface orientation or illumination changes (Hansen & Gegenfurtner, 2009; Kingdom, 2003).

As an example, Figure 1 (left), shows a color image of tomatoes in a basket, from Olmos and Kingdom (2004b). This color image can be converted into L, M, and S cone quantal catches (Olmos & Kingdom, 2004a), which can then be combined into a luminance (L+M) channel and a red-green opponent (L-M) channel. Edges in these images can be found by taking the norm of the gradient of image intensities  $I(x, y)$ ,

given by  $\sqrt{(\partial I(x, y)/\partial x)^2 + (\partial I(x, y)/\partial y)^2}$ . The luminance edges shown in Figure 1 (middle) indicate both object boundaries and illumination boundaries: For example, both the edge of the basket and the shadow cast by the basket's lip are marked here. However, the red-green edges (Figure 1, right) only pick up the boundary of the basket and ignore the lip's shadow within the basket boundary. By using the output of luminance and color edge detectors, the shadow edge of the lip could easily be disambiguated from the object boundary just above it.

Another hint at the usefulness of color edge detectors comes from neural networks. “Alexnet” (Krizhevsky, Sutskever, & Hinton, 2012) was developed to recognize objects in color images. The first layer of Alexnet is a set of convolutional filters that sparsely represent the image. These filters are shown in Figure 2. Many of the filters are readily interpretable as color edge detectors.

For these reasons, a system for processing shape and form in color vision would appear to be very useful for exploiting the chromatic object boundary information in the visual scene. Its existence, however, was initially doubted, because of the psychophysical low-pass, low acuity color contrast sensitivity function (CSF; Kim, Reynaud, Hess, & Mullen, 2017; Mullen, 1985), which is not indicative of edge detectors, and physiological reports of a lack of orientation tuning for color in the

Citation: McIlhagga, W., & Mullen, K. T. (2018). Evidence for chromatic edge detectors in human vision using classification images. *Journal of Vision*, 18(9):8, 1–17, <https://doi.org/10.1167/18.9.8>.

<https://doi.org/10.1167/18.9.8>

Received March 9, 2018; published September 7, 2018

ISSN 1534-7362 Copyright 2018 The Authors



This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

Downloaded From: <https://jov.arvojournals.org/pdfaccess.ashx?url=/data/journals/jov/937491/> on 09/27/2018

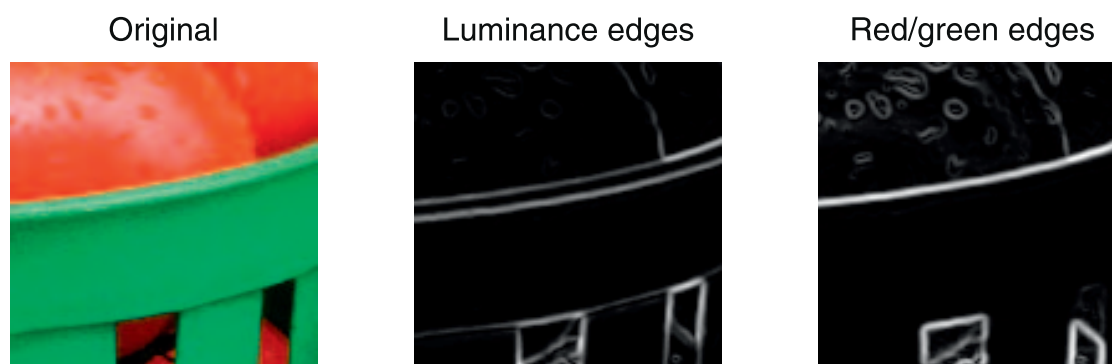


Figure 1. The left hand image shows a color photograph of tomatoes in a green plastic basket, downloaded from Olmos and Kingdom (2004b). The luminance edges (center image) mark discontinuities in the intensity of the luminance of the image, and the red/green edges (right image) show discontinuities in the red/green balance of the image. A more detailed analysis is presented in Johnson and Mullen (2016).

primate visual system (Livingstone & Hubel, 1984, 1987). Subsequent psychophysical studies, however, suggested that bandpass spatial filtering with broadly similar bandwidths for color and achromatic contrast underlies the overall low-pass shape of the CSF (Bradley, Switkes, & De Valois, 1988; Humanski & Wilson, 1992; Losada & Mullen, 1994, 1995; Mullen & Losada, 1999). Furthermore, psychophysical studies have demonstrated orientation tuned responses in color vision (Beaudot & Mullen, 2005; Bradley et al., 1988; Humanski & Wilson, 1993; Reisbeck & Gegenfurtner, 1998; Vimal, 1997; Webster, Switkes, & Valois, 1990; Wuerger & Morgan, 1999), although this may be lost at very low spatial frequencies (Gheiratmand, Meese, & Mullen, 2013; Gheiratmand & Mullen, 2014). Thus, the presence of both spatial frequency and orientation tuning neural responses, the prerequisites for edge detection, suggest that there may be a system for chromatic edge detection.

Here, we use classification images (Ahumada, 1996; Beard & Ahumada, 1998; Murray, 2011) to look for evidence of chromatic edge detectors in the human visual system for both L/M (red-green) and S cone (blue-yellow) cone opponent responses. In our experiments, observers had to detect a luminance edge in luminance noise, a red/green isoluminant edge in red/green isoluminant noise, or a blue/yellow isoluminant edge in blue/yellow isoluminant noise. For luminance edges, we find that the classification image for edge detection is similar in shape to a derivative of Gaussian filter, which is optimal for Gaussian edges (Lindeberg, 1998; McIlhagga, 2011), and is consistent with previous psychophysical evidence for edge detectors (Shapley & Tolhurst, 1973). For red/green and blue/yellow isoluminant edges, the classification image is also like a derivative of Gaussian filter, but somewhat wider than those for the luminance edges. If the classification images discovered for luminance edges are evidence for



Figure 2. The convolutional filters from the first layer of the Alexnet neural network (Krizhevsky et al., 2012). Each filter is 11 by 11 pixels. The neural net was implemented on 2 GPUs, which did not communicate until higher layers. (A GPU is a graphics processing unit and is frequently used to speed neural net training.) The top 48 filters are from GPU1 and the bottom 48 from GPU2. The filters on one GPU tend to specialize in encoding luminance changes, and the filters on the other GPU tend to specialize in color. Reprinted with permission from Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet Classification with Deep Convolutional Neural Networks. In F. Pereira, C. J. C. Burges, L. Bottou, & K. Q. Weinberger (Eds.), *Advances in Neural Information Processing Systems 25* (pp. 1097–1105). Curran Associates, Inc.

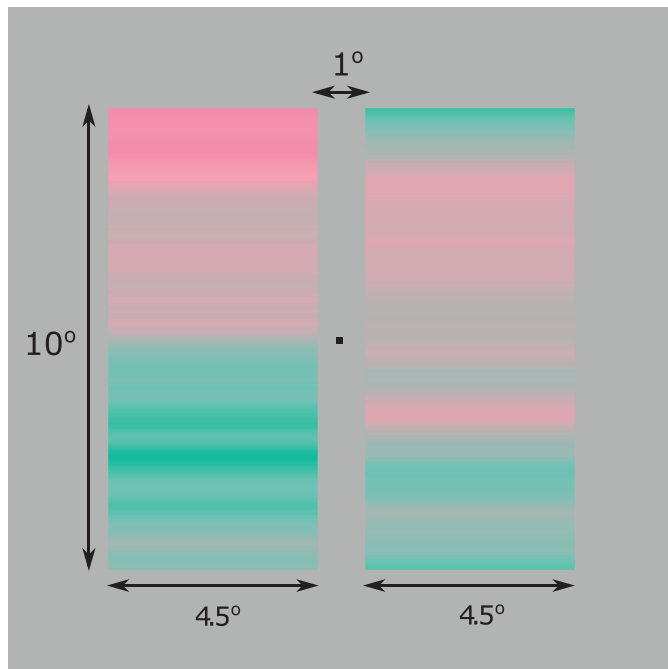


Figure 3. An example of the red-green stimuli used in the experiments. The left-hand image in this case contains a Gaussian blurred edge and filtered Brown noise whereas the right-hand image only contains filtered Brown noise. In this example, the edge is a hue change from red above to green below. Not all edges involved a change in hue. The small fixation dot in the middle was continuously visible, and the edge was always horizontally aligned with it. The gray surround has been cropped.

luminance edge detectors, then our classification image results demonstrate that the red/green and blue/yellow chromatic channels may also contain edge detectors.

## Methods

### Brief outline

Experiments were conducted at two sites, Bradford University, UK, and McGill University, Canada. A two-alternative forced-choice task was used in which the observer was shown two images side by side, one with a horizontal edge, and one without (Figure 3). The observer's task was to indicate by pressing a mouse button which image contained the edge. Both images contained filtered Brown noise, used in the classification image analysis. There were three kinds of edge/noise stimuli: luminance, red-green isoluminant, and blue-yellow isoluminant. Edges were always shown with a fixed polarity (dark above, light below for the luminance edge; green above and red below for the red/green edge; yellow above, blue below for the blue/

yellow edge) and they were always presented at the vertical center of the stimulus.

### Observers

There were five observers in the study. All had normal or corrected-to-normal visual acuity and normal color vision as assessed with the Farnsworth-Munsell 100 Hue test. The experiments were performed in accordance with the Declaration of Helsinki and approved by the institutional ethics committee of the Research Institute of McGill University Health Centre and the Ethics procedure at Bradford University School of Life Sciences. All participants signed an information consent form.

### Color space

Stimulus color and contrast are described in cone-contrast space (Cole, Hine, & McIlhagga, 1993; Stromeyer, Cole, & Kronauer, 1985), defined as follows. If  $(L_b, M_b, S_b)$  is the vector of cone quantal catches of the  $L$ ,  $M$ , and  $S$  cone photoreceptors on the background, [computed following Cole and Hine (1992)], and  $(L_x, M_x, S_x)$  is the vector of cone quantal catches at some point  $x$  on the stimulus, then the cone contrasts at that point are given by the vector

$$\left( \frac{L_x - L_b}{L_b}, \frac{M_x - M_b}{M_b}, \frac{S_x - S_b}{S_b} \right)$$

A stimulus color is specified by a vector  $(l, m, s)$ , normalized to a unit length ( $l^2 + m^2 + s^2 = 1$ ). The stimulus contrast at point  $x$  is specified by a number  $C(x)$ . The cone contrast of the stimulus at location  $x$  is the product of these,

$$\left( \frac{L_x - L_b}{L_b}, \frac{M_x - M_b}{M_b}, \frac{S_x - S_b}{S_b} \right) = C(x) \times (l, m, s) \quad (1)$$

The direction  $(l, m, s)$  of the cone contrast vector is the same at every point, but the length  $C(x)$  varies. Stimulus color vectors were designed to isolate the luminance (achromatic), red-green, or blue-yellow postreceptoral cone opponent mechanisms (Cole et al., 1993; Krauskopf, Williams, & Heeley, 1982; Sankeralli & Mullen, 1996). The luminance stimulus has a direction of  $(l, m, s) = (1, 1, 1)$ , the blue-yellow direction is the  $S$  cone axis  $(l, m, s) = (0, 0, 1)$  and the red-green isoluminant stimulus has a direction  $(l, m, 0)$  where the values of  $l$  and  $m$  are determined individually for each subject, using a minimum perceived motion technique (Anstis & Cavanagh, 1983). Each subject varied the ratio of  $L$  and  $M$  cone contrast by method of adjustment to find a minimum in the perceived motion



of horizontal Gabor with a drifting sinewave carrier (3 Hz, 1 c/°). An average of about 10 measurements was taken as the individual's isoluminant point.

## Stimuli

Stimuli consisted of two images presented side by side on a gray background (Figure 3). Each image was 10° high and 4.5° wide, and they were separated horizontally by 1°. One image contained an edge and smoothed Brown noise and the other image only contained smoothed Brown noise. A small fixation mark was provided at the center of the screen. The edge was generated by blurring a step edge with a Gaussian filter, using filter scales of  $\sigma_e = 0.1125^\circ$ ,  $0.225^\circ$ , or  $0.45^\circ$ . The Brown noise in the image was also horizontal. Brown noise is the integral of white noise, so the Brown noise in these experiments was created by generating independent normally distributed white noise samples for each scan line in the image, then computing their cumulative sum. The resultant Brown noise was then blurred by a Gaussian filter with a scale of  $\sigma_b = 0.1125^\circ$ , with the aim of reducing artifacts generated by chromatic aberration (at 3 c/°, the noise amplitude is reduced by 89%). The Brown noise sample was shifted to have a mean of zero. Brown noise contrast can be specified by the drift rate per degree: if  $b(x)$  is the brown noise as a function of position  $x$  in degrees, the drift rate is  $\sqrt{E((b(x) - b(x+1))^2)}$ . The drift rate was 0.075, 0.02, and 0.16 per degree for luminance, red/green, and blue/yellow stimuli respectively. These were chosen to produce reasonably high contrast stimuli at threshold within the available monitor gamut.

The use of Brown noise differs from most classification image experiments, which typically use white noise. However, if we used white noise, then stimulus areas far from the edge would provide cues to the existence of the edge. For example, an average increase in redness at the bottom of the stimulus would provide a cue for a red/green edge, even if the edge itself was not seen. For this reason, we used Brown noise rather than white noise. Brown noise is also the more appropriate form of noise for mapping localized edge detector mechanisms, because the optimal edge detector filter (using the Canny (1986) criteria) is only spatially localized in the presence of Brown noise (McIlhagga, 2011). Brown noise is also more ecologically relevant, since the power spectrum of Brown noise ( $1/f^2$ ) is nearly the same as the power spectrum of natural images (Burton & Moorhead, 1987; Field, 1987; for a survey, see Billock, 2000) and the adult visual system may be matched to this sort of power spectrum (Billock, 2000).

## Experimental procedure

A staircase method was used to adjust the contrast of the edge. In each trial, the observer chose which image, left or right, they thought contained the edge by pressing the left or right mouse button. The contrast of the edge was adjusted in equal log-steps by a one-up two-down staircase (Levitt, 1971). Audio feedback was given after each response. Stimuli were displayed for 500 ms, using a Gaussian temporal window with scale  $\sigma = 125$  ms. After the observer responded, there was a 1 s delay before the next stimulus appeared. Responses were collected from observers in blocks of 150 trials. Between nine and 27 blocks of data were collected for each combination of subject, edge width, and color. Edge width and color direction were kept constant within each block.

## Calibration and apparatus

Stimuli were displayed on cathode ray tube (CRT) monitors driven by a Bits++ device in Colour++ mode (Sony Multiscan E450 CRT at Bradford, Mitsubishi Diamond Pro CRT at McGill; Cambridge Research Systems Ltd., Kent, UK). In this mode, adjacent eight-bit pixels in the frame buffer are paired to yield 16 bits per pixel for each CRT gun, and the 14 most significant bits are passed to a digital-to-analog converter. Stimuli were calculated and displayed by MATLAB (MATLAB Release 2007b, MathWorks, Natick, MA), using the Psychophysics Toolbox (Brainard, 1997; Kleiner, Brainard, & Pelli, 2007; Pelli, 1997). The gamma of the monitors was corrected using an OptiCal photometer (at McGill) or a ColorCal meter (at Bradford), both from Cambridge Research Systems. Display resolution and viewing distances were slightly different in both locations, which meant that the stimuli occupied a different number of scanlines for the same angular size. All stimuli were therefore interpolated to the same resolution before classification image analysis. The spectral radiances of the red, green, and blue phosphors of the monitor were calibrated using a PR-645 SpectraScan spectroradiometer at McGill and a PR-650 Spectrascan at Bradford (both Photo Research Inc., Chatsworth, CA).

## Classification images

Classification images are usually estimated by weighted sums (Beard & Ahumada, 1998; Murray, 2011; Murray, Bennett, & Sekuler, 2002). Here, they are estimated using logistic regression (Knoblauch & Maloney, 2008; McCullagh & Nelder, 1989; Mineault, Barthelmé, & Pack, 2009). In outline, the classification

image procedure works as follows. On the  $i$ -th trial of an experiment, the edge image has contrast at position  $x$  given by

$$(C_i(x) + n_i(x)) \times (l, m, s)$$

where  $C_i(x)$  is the profile of the Gaussian edge at position  $x$  on the  $i$ -th trial,  $n_i(x)$  is the added Brown noise on trial  $i$ , and  $(l, m, s)$  is the cone-contrast direction used (see Equation 1). The non-edge image has contrast given by

$$u_i(x) \times (l, m, s)$$

where  $u_i(x)$  is the Brown noise at position  $x$  in this stimulus on trial  $i$ .

The data can be analyzed by assuming that the human observer is linear. A linear observer makes their decision by computing a weighted sum of the contrasts in each stimulus, and then taking the difference. The set of weights used by a linear observer is called a template. With a color stimulus, there are three templates, one for each cone class, which we will call  $t_x^L$ ,  $t_x^M$ , and  $t_x^S$ , where the subscript gives the position ( $x$ ) and the superscript ( $L$ ,  $M$ , or  $S$ ) says which cone type the template is applied to.

Using this template, the weighted sum of contrasts for the edge stimulus is

$$\begin{aligned} a_i &= \sum_x t_x^L \times l(C_i(x) + n_i(x)) \\ &\quad + \sum_x t_x^M \times m(C_i(x) + n_i(x)) \\ &\quad + \sum_x t_x^S \times s(C_i(x) + n_i(x)) \\ &= \sum_x t_x(C_i(x) + n_i(x)) \end{aligned}$$

where  $t_x = (t_x^L l + t_x^M m + t_x^S s)$  is the dot product of the template  $(t_x^L, t_x^M, t_x^S)$  and the cone contrast direction  $(l, m, s)$ . The weighted sum of contrasts for the non-edge stimulus is

$$\begin{aligned} b_i &= \sum_x t_x^L \times l u_i(x) + \sum_x t_x^M \times m u_i(x) \\ &\quad + \sum_x t_x^S \times s u_i(x) \\ &= \sum_x t_x u_i(x) \end{aligned}$$

Thus the decision variable on trial  $i$  is:

$$d_i = a_i - b_i = \sum_x t_x(C_i(x) + n_i(x) - u_i(x)) \quad (2)$$

Because the cone contrast direction is fixed in an experiment, it is not possible to estimate the individual components of the template, only the dot product  $t_x$ , and this is what is shown in the figures in the Results section.

The observer's decision is based on the value of the decision variable. If there was no internal noise, the observer would be correct if  $d_i > 0$  and incorrect otherwise. We assume, however, that there *is* internal noise added to the decision variable, and so the observer is correct if  $d_i + e_i > 0$ , where  $e_i$  is an additional internal noise term, and incorrect otherwise. For simplicity, we assume that  $e_i$  has a standard logistic distribution, so the probability that the observer is correct is given by

$$pr(d_i + e_i > 0) = (1 + \exp(-d_i))^{-1} = p_i \quad (3)$$

Unfortunately, the observer's decision variable  $d_i$  cannot be observed directly, but we record whether they are correct or incorrect. Let  $r_i$  be 1 if the observer is correct on the  $i$ -th trial, and 0 otherwise. The log-likelihood  $L$  of the observer's responses  $r_i$ , given the above probabilities  $p_i$ , is given by the sum of the log-likelihoods of each individual trial, each of which follows a Bernoulli distribution (i.e., a Binomial distribution with  $n = 1$ ):

$$L = \sum_i r_i \log p_i + (1 - r_i) \log(1 - p_i) \quad (4)$$

Since the log-likelihood implicitly depends upon the template values  $t_x$ , those values can be estimated by maximizing the log-likelihood  $L$ . A classification image is simply an estimate of the template  $t_x$ . Full details of the estimation procedure are given in Appendix A.

Classification images estimate a linear observer that fits human responses as closely as possible. It is a separate question whether this fit is statistically close enough to accept. The standard method for assessing goodness of fit for logistic regression (Hosmer & Lemeshow, 1980) has substantial flaws (Kuss, 2002). We used a Monte Carlo test to assess goodness of fit. If the model probabilities  $p_i$  are correct, we can simulate new observations. Let  $r_i^{(1)}$  be the  $i$ -th simulated response, where  $r_i^{(1)} \sim \text{Bernoulli}(p_i)$ . We can calculate a log-likelihood from the simulated responses, namely  $L^{(1)} = \sum_i r_i^{(1)} \log(p_i) + (1 - r_i^{(1)}) \log(1 - p_i)$ . By repeating this process many times, we create a set of simulated likelihoods  $L^{(1)}, L^{(2)}, L^{(3)}, \dots$ , which is the distribution of likelihoods under the hypothesis that the fitted probabilities  $p_i$  are correct. These can then be compared to the actual likelihood  $L$  computed from the actual responses. Now define the  $p$  value for the classification image as the fraction of simulated likelihoods which are less than  $L$ . If this  $p$  value is extreme (say less than 2.5% or greater than 97.5%), the observed data is unlikely to have come from the model probabilities  $p_i$ , and so we would be inclined to doubt the validity of the model. If it is not extreme, we would be inclined to accept it.

## Results

### Classification images

We estimated 45 classification images for all combinations of the five observers, three color directions, and three edge widths. As an example, Figure 4 shows the classification images obtained for all observers at edge scale  $\sigma_e = 0.225^\circ$ . (The plots for other edge scales are available in the Supplementary Material.) The columns show the classification images for the luminance, red/green, and blue/yellow color directions. The rows show the classification images for each observer A, B, C, D, and E (A and E are the authors). Most of the classification images are similar to a derivative of Gaussian filter, given by  $f(x) = \alpha(x - c)^2 \exp(-(x - c)^2 / (2\sigma_f^2))$ , where  $\sigma_f$  is the Gaussian filter scale, and  $c$  and  $\alpha$  are the center and amplitude. The best fit filters are shown by the broad pale lines in Figure 4. Detection thresholds were calculated by fitting a Weibull function to the frequency-of-seeing data. These thresholds, divided by the Brown noise drift rate for each color direction, are shown as gray bars in each panel. The length of the bar is proportional to the threshold. Although the thresholds vary between observers, within each observer there tends to be only a small variation across the different color directions.

These classification images do not, however, completely resolve the mechanism underlying edge detection; rather, they are the best linear approximation to whatever mechanism is actually used by our observers to detect the edges. The simplest interpretation of our results is that observers apply an edge detection filter, like those shown in Figure 4, to a contrast signal, whether that contrast signal is luminance ( $L + M$ ), red-green ( $L - M$ ) or blue-yellow ( $S - (L + M)/2$ ). However, it is also possible that in the color conditions the observers are looking for a change in hue from, say, red to green, rather than just a change in contrast. This may be because the color channels are opponent and encoded in separate red and green or blue and yellow processes.

A subset of our data is relevant to this question. The brown noise sometimes creates stimulus pairs where the contrast signal ( $C(x)$  in Equation 1) does not cross zero within the central  $2^\circ$  of each stimulus in the pair. This happened in about one-fourth of trials and was more likely with a lower contrast edge. In the stimulus shown in Figure 3, left panel, the contrast does cross zero at the edge location when it goes from red to green. In the red-green condition, the no-zero crossing condition means that the cone contrast signal  $L - M$  is either entirely positive (red) or entirely negative (green) within the central  $2^\circ$ ; in

the blue-yellow condition,  $S > 0$  or  $S < 0$  within the central  $2^\circ$ . For the color conditions, no zero-crossings mean the edge is marked by a change in saturation, for example from pale red to a more intense red. For the luminance condition, no zero-crossings means the stimulus luminance was either entirely above or entirely below the average luminance within the central  $2^\circ$ .

When restricted to just this set of trials, the classification images produced are shown in Figure 5. As expected, the luminance classification images do not change much at all. For the color conditions, observers A, C, and D have essentially the same classification images as in the full set of trials (Figure 4). Thus, they appear to be detecting edges by looking for any differences color *contrast*, rather than specifically hue changes. Observer B did not collect very much data, so their classification images are basically noise with this even smaller dataset. Observer E is more interesting. Although their classification image for luminance in Figure 5 is the same as in Figure 4, their classification images for the color directions are mostly random. That is, they seem unable to detect color edges that do not involve a change in hue. This could be due to a low-level process, such as them being unable to see saturation changes; or a high-level cause, such as them detecting edges like the other observers, but consciously rejecting those that do not have a hue change.

### Goodness of fit

The  $p$  values obtained for the classification images in Figure 4, using the Monte Carlo goodness-of-fit test with 64,000 repetitions, are shown in Table 1.

None of the  $p$  values in Table 1 are below 0.025 or greater than 0.975, so classification image model provides an acceptable fit to the data. Under the hypothesis of the model being correct, we would expect  $p$  values to be uniformly distributed between 0 and 1. This is not the case here—all of the  $p$  values are higher than 0.7. This is a consequence of using a penalized logistic regression to estimate the classification images. The penalty reduces the magnitude of the template values, which in turn moves the decision variable closer to 0. This then makes the fitted probabilities closer to 0.5 than they would be without the penalty. This shift towards 0.5 causes the increase in  $p$  values.

### Detection thresholds

Since the noise level in these experiments is quite high, detection thresholds are influenced more by the level of added noise than the intrinsic sensitivity of the

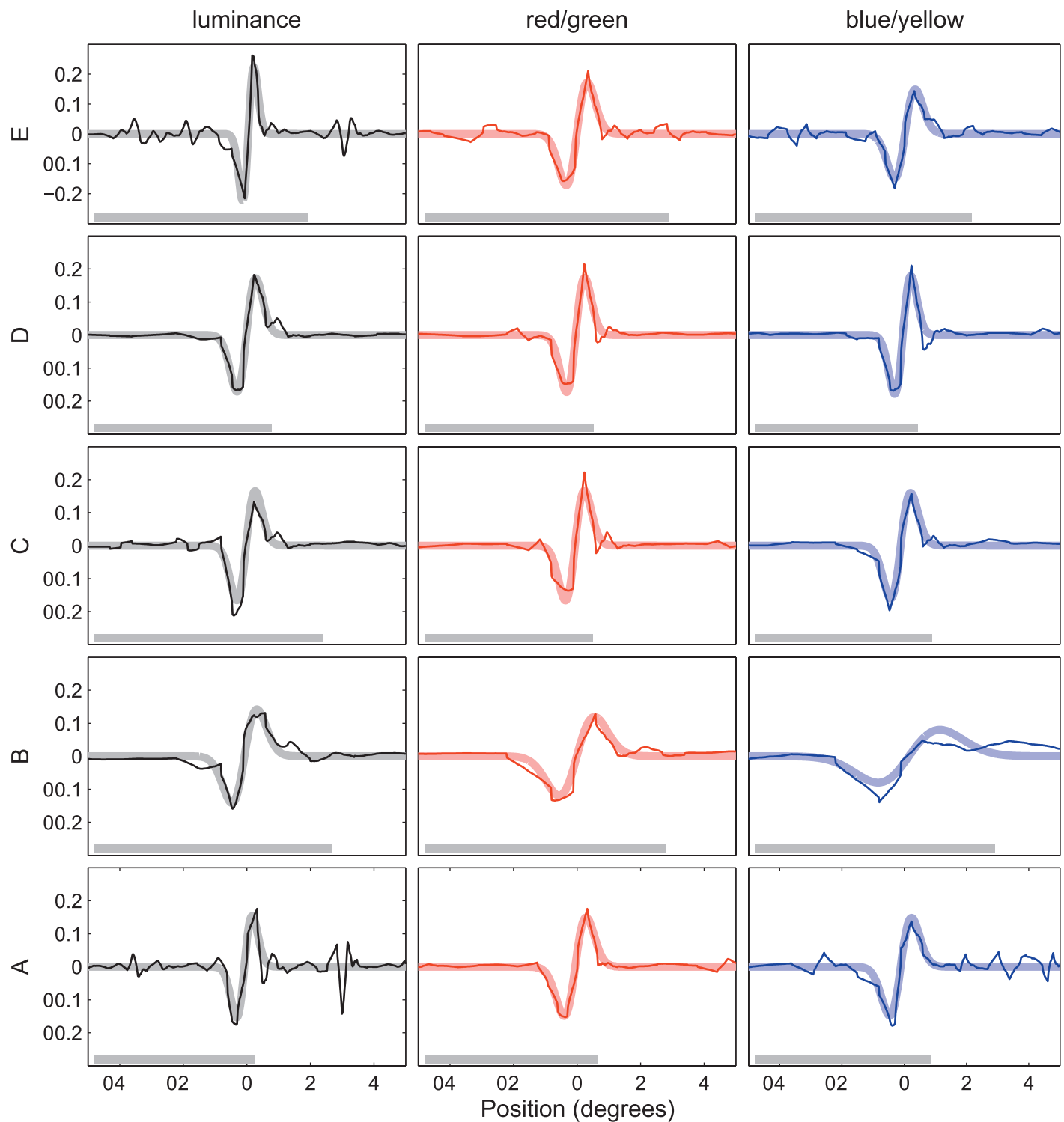


Figure 4. Classification images for five observers (A, B, C, D, and E), and three color directions (luminance, red/green, and blue/yellow) for the edge with scale  $\sigma_e = 0.225^\circ$ . The classification images in this figure have been normalized to the same power to make comparison of their shapes easier. The broad pale lines show the best-fitting derivative of Gaussian filter. The length of the gray bars in the lower left of each panel are proportional to the cone-contrast threshold for detecting the edge, divided by the drift rate of the noise for that condition. The classification images for the other edge widths are given in the Supplementary Material.



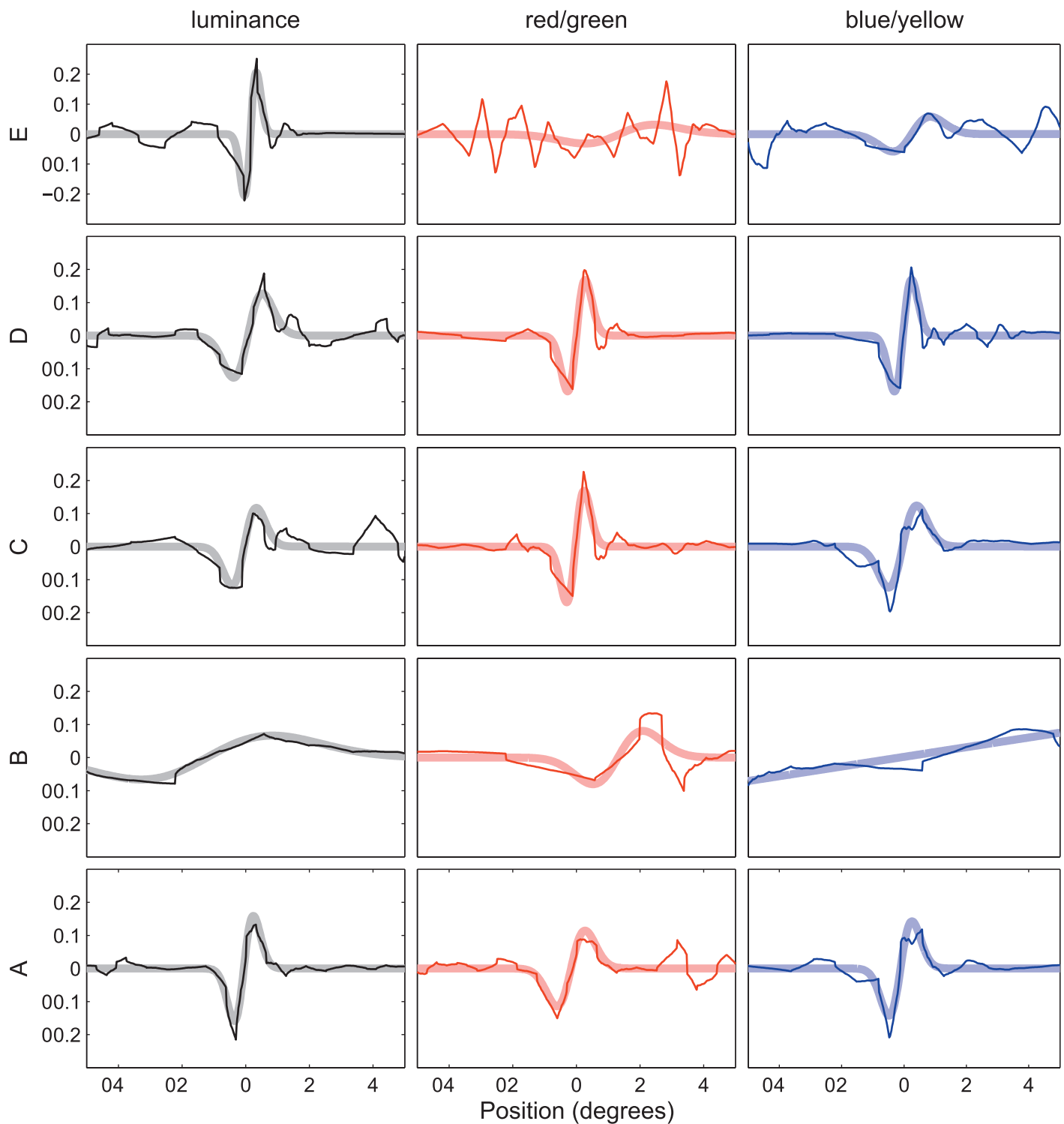


Figure 5. Classification images for the stimulus conditions of Figure 4, but estimated from trials where the contrast does not cross zero within the central  $2^\circ$  of both stimuli. For the color conditions, this implies that these stimuli do not have a hue change (red to green, or blue to yellow) in the central  $2^\circ$ . Observer B did not collect much data overall, so with even less data here their classification images are not meaningful. The results from the other observers are discussed in the text.

edge-detection system. Thus, the best measure of threshold is to scale the raw threshold by the level of the noise. Figure 6 shows the edge detection thresholds, divided by the Brown noise drift rate, averaged across

all subjects. The differences between the luminance, red/green, and blue/yellow thresholds are minor, suggesting that they are the result of similar mechanisms.



SUBJ	$\sigma_e = 0.1125^\circ$			$\sigma_e = 0.225^\circ$			$\sigma_e = 0.45^\circ$		
	lum	r/g	b/y	lum	r/g	b/y	lum	r/g	b/y
E	0.81	0.83	0.88	0.79	0.83	0.8	0.8	0.85	0.81
D	0.95	0.9	0.73	0.97	0.95	0.86	0.92	0.92	0.84
C	0.9	0.89	0.97	0.92	0.95	0.92	0.81	0.93	0.84
B	0.85	0.9	0.88	0.85	0.92	0.87	0.81	0.83	0.85
A	0.96	0.96	0.92	0.82	0.88	0.89	0.88	0.83	0.85

Table 1. Goodness-of-fit  $p$  values for the classification images for all conditions.  $P$  values less than 0.025 or greater than 0.975 would indicate a possible failure of the model.

## Classification image widths

The classification images for all observers show a similar shape within any stimulus condition, so it is reasonable to summarize our data by averaging across observers. Figure 7 shows the classification images averaged across the five observers, for all three edge scales and all color conditions. Classification images have been multiplied by the Brown noise drift rate for each color direction to equalize their amplitude. The gray bars in each panel cover the interval  $x = \pm 2\sigma_e$  for each edge scale  $\sigma_e$ . The average classification images in Figure 7 are like derivatives of Gaussian filters, and their width roughly follows the edge scale.

It is difficult to specify the width of the classification images precisely, given the random bumps in them. To get a robust measure of classification image width, we fitted derivative of Gaussian filters

$f(x) = \alpha(x - c)^2 \exp(-(x - c)^2 / (2\sigma_f^2))$  to all 45 classifi-

cation images that we estimated. The fit had three free parameters:  $\sigma_f$  which is the Gaussian filter scale, and  $c$  and  $\alpha$ , which are center and amplitude. The filter scales  $\sigma_f$  were then averaged across observers. These averages and standard errors are plotted in Figure 8. The filter scales for the luminance conditions are most similar to the edge scales (shown as a gray line). The filter scales for the red/green and blue/yellow conditions suggest that these channels may have a minimum scale of about  $\sigma_f = 0.32^\circ$  for red/green and  $0.4^\circ$  for blue/yellow conditions.

## Discussion

We have estimated classification images for an edge detection task using three different edge widths and three directions in color space (achromatic, red/green isoluminant, and the S cone isolating blue/yellow direction). The classification images for edge detection in all three color directions are like derivative of Gaussian filters and are consistent with the presence of edge detection filters in the visual system. However, some differences are apparent. For the luminance edge, the filter scale, as measured by the best fit derivative of Gaussian, is close to the edge scale for all three edges scales tested. For both chromatic edges, however, the filter scale is close to the edge scale only for the most blurred edge and as the edge narrows the filter scale remains relatively broad. And in one observer, edges were not detected when there was no change in hue.

Edge detection filters are naturally represented in the space domain. However, filters in spatial vision are frequently displayed in the Fourier domain. In Figure 9 we have computed the smoothed amplitude spectrum for the average classification images in Figure 7. These spectra have been plotted, together with estimates of the human CSF. Based on these graphs, it is reasonable to identify these edge detectors with broad spatial frequency channels.

The peak spatial frequency for the luminance filters is only about  $2 \text{ c}/^\circ$ , somewhat lower than the peak of the human CSF. It is probable that the filters we have

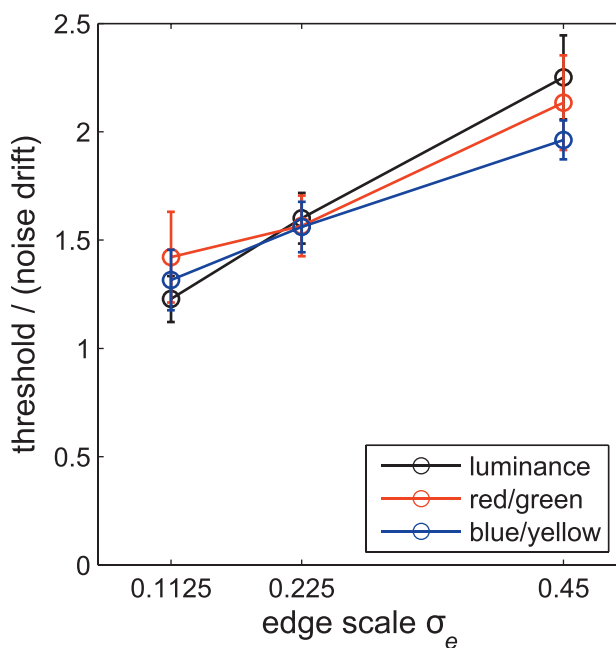


Figure 6. Edge detection thresholds, divided by the Brown-noise drift rate, for the three color directions and edge scales used in this study. Vertical bars show standard errors.

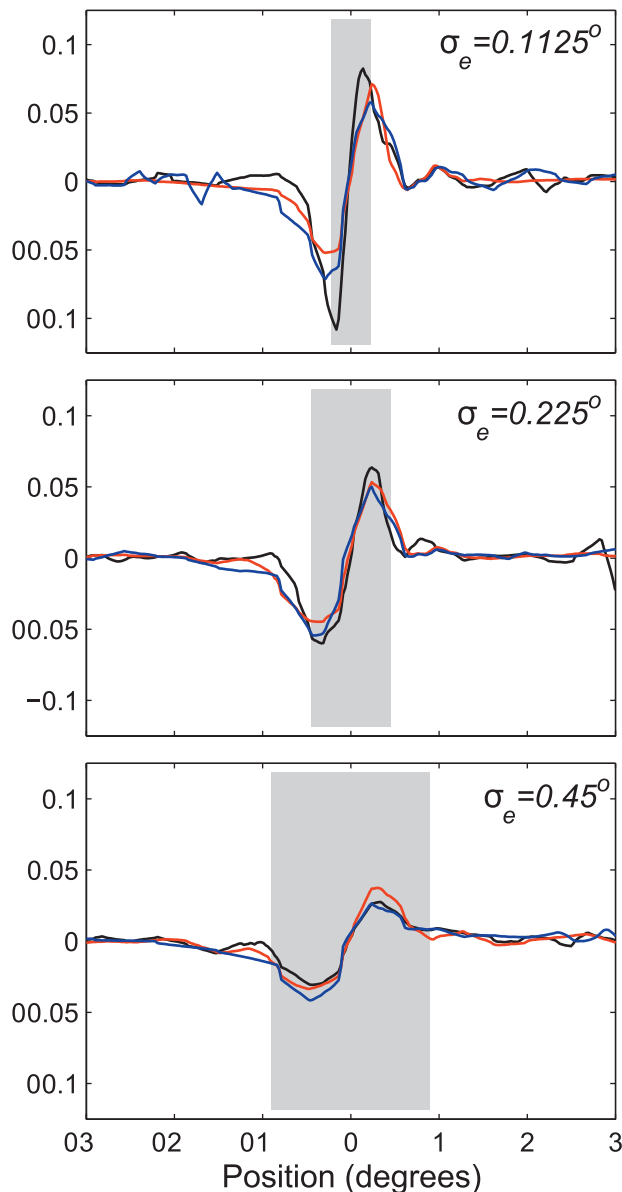


Figure 7. In each panel, the average classification image for the luminance color direction is black, for the red/green color direction is red, and for the blue/yellow color direction is blue. The vertical gray bars have a width of  $4\sigma_e$  for the edge in question (e.g., for the top panel, the edge scale  $\sigma_e$  is  $0.1125^\circ$  and the gray bar has a width of  $0.45^\circ$ ). Most of the edge lies within this bar. Classification images were averaged across subjects, then scaled by the brown noise drift rate.

found represent some of the broadest spatial frequency channels in the luminance system. The picture is somewhat different for the chromatic channels. Here, the spatial frequency profile of the filters is towards the higher end of the human CSF. Thus, the filters we have found are not the lowest spatial frequency channels for red/green or blue/yellow systems. Instead, there must be even lower spatial-frequency channels in the chromatic system; these could be simply “blob”

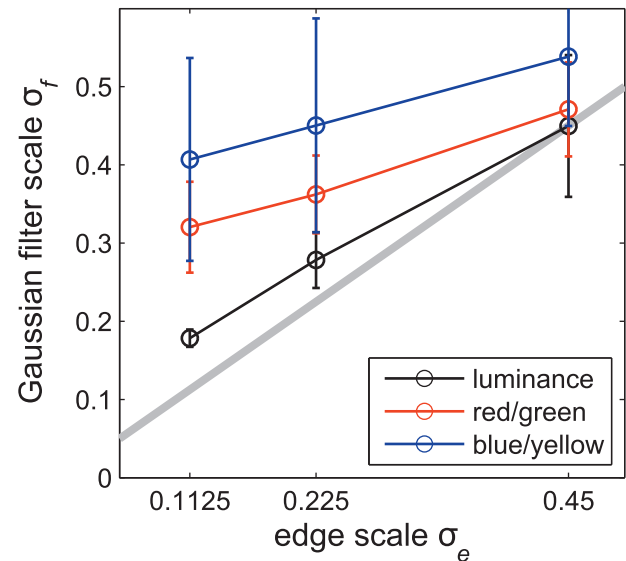


Figure 8. Average derivative of Gaussian filter scales  $\sigma_f$  in degrees are plotted as a function of the edge scale  $\sigma_e$  (in degrees) averaged across subjects ( $N = 5$ ). The gray line shows where the filter scale equals the edge scale.

detectors with a low-pass rather than a band-pass characteristic. There is physiological and psychophysical evidence for low pass non-oriented chromatic detectors with high contrast sensitivity in primate vision (Gheiratmand et al., 2013; Schluppeck & Engel, 2002; Shapley & Hawken, 2011).

## Optimal edge detectors

The idea of an optimal edge detector was introduced by Canny (1986). He suggested that an optimal edge detector should have a good signal-to-noise ratio and localize the edge well. This work was extended by McIlhagga (2011) who showed that the optimal edge detector in white noise was infinitely wide, but had a finite width in Brown noise (which has the same  $1/f^2$  power spectrum as real images). The optimal detector of a Gaussian-blurred edge in brown noise is approximately a derivative of Gaussian filter, whose scale is matched to the edge scale (McIlhagga, 2011). Such filters have been suggested as a component of human edge perception (Georgeson, May, Freeman, & Hesse, 2007) and blur discrimination (McIlhagga & May, 2012).

However, the experiments here used *filtered* brown noise. The filtering was necessary to reduce the effects of chromatic aberration, but it changes the optimal detector. Using just the signal-to-noise criterion (localization being complicated and having minor effect on the filter), the optimal detector of a Gaussian blurred edge, with scale  $\sigma_e$ , in blurred Gaussian Brown noise, with blur scale  $\sigma_b$ , is a derivative of Gaussian

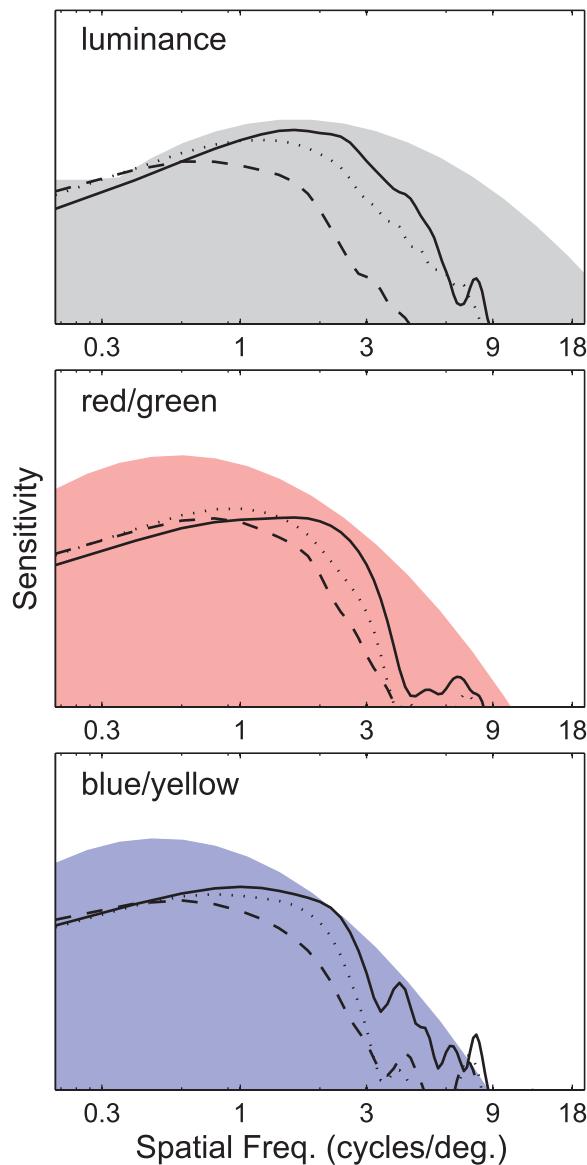


Figure 9. Smoothed amplitude spectra of the classification images plotted in Figure 7. Each panel is a different color direction. Within each panel, the solid line is the amplitude spectrum for the classification image at edge scale  $\sigma_e = 0.1125^\circ$ , the dotted line for  $0.225^\circ$ , and the dashed line  $0.45^\circ$ . The shaded areas plot the human CSF, scaled by eye, for comparison, from Kim et al., (2017).

filter with scale  $\sigma_f = \sqrt{\sigma_e^2 - 2\sigma_b^2}$ , provided the expression under the square root is not negative (see Appendix B). Thus the optimal detector in filtered brown noise is narrower than the edge scale  $\sigma_e$ . However, the scale of the classification image filters here is *larger* than the edge scale, so they are not optimal. There are a couple of possible explanations for this suboptimality.

First, if the observers had any spatial uncertainty about the location of the edge, this would have the effect of widening the classification image. The

classification image is only equal to an edge detection filter if the output of that filter always peaks at the edge location. If not, then the classification image is the average of the filter over the locations where it peaked on separate trials. This has the effect of blurring out the classification image, and increasing its apparent width. Although this might account for some of the discrepancy between the scale of the classification image and the optimal scale, it cannot account for all of it.

Secondly, the filter scale is less important for the wider edges than for the narrowest edge. The signal-to-noise ratio for a derivative of Gaussian with scale  $\sigma_f$  detecting an edge with scale  $\sigma_e$  in Brown noise with blur  $\sigma_b$  is given by

$$SNR \propto \sqrt{\frac{\sigma_f^2 + \sigma_b^2}{\sigma_f^2 + \sigma_e^2}}$$

(from Appendix B). The left panel of Figure 10 plots the signal-to-noise ratio, as a function of  $\sigma_f$ , for the three edge scales used in this study. The circles show the average filter scales obtained from Figure 8. For the two widest edges, the optimum is broad and the classification images are hardly suboptimal at all.

Once you know the theoretical signal-to-noise ratio of a filter, you can invert it to get the theoretical detection threshold for that filter. We calculated theoretical thresholds for the best-fit derivative of Gaussian filters with widths taken from Figure 8. These are shown in Figure 10, right panel. Although the ordering of the thresholds is incorrect with respect to color channel, the shape of the threshold curves is broadly the same as those plotted in Figure 6. The gray curve in the left panel of Figure 10 shows the threshold that would be obtained using the optimal width filter. Apart from the narrowest edge, there are only small differences between the optimal threshold and the thresholds obtained using our derivative of Gaussian filters.

## Classification images and receptive fields

Although there is no necessary link between classification images and neural receptive fields, it is interesting to see how consistent our classification images are with them. The peak spatial frequency of a derivative of Gaussian with standard deviation  $\sigma$  is  $1/\sqrt{2\pi}\sigma$ . Thus, the luminance filters have peak spatial frequencies of 1.4, 0.9, and 0.5 c/°, for the three edge stimuli of widths  $\sigma_e = 0.1125^\circ$ ,  $0.225^\circ$ , and  $0.45^\circ$ , respectively. This is within the range found in electrophysiological studies in primates. Of course, because of eye movements, spatial uncertainty, horizontal summation, and the range of filters with different peak sensitivity functions, there is no possibility that the classification images we have found are of a single V1 receptive field. However, they could be like the average

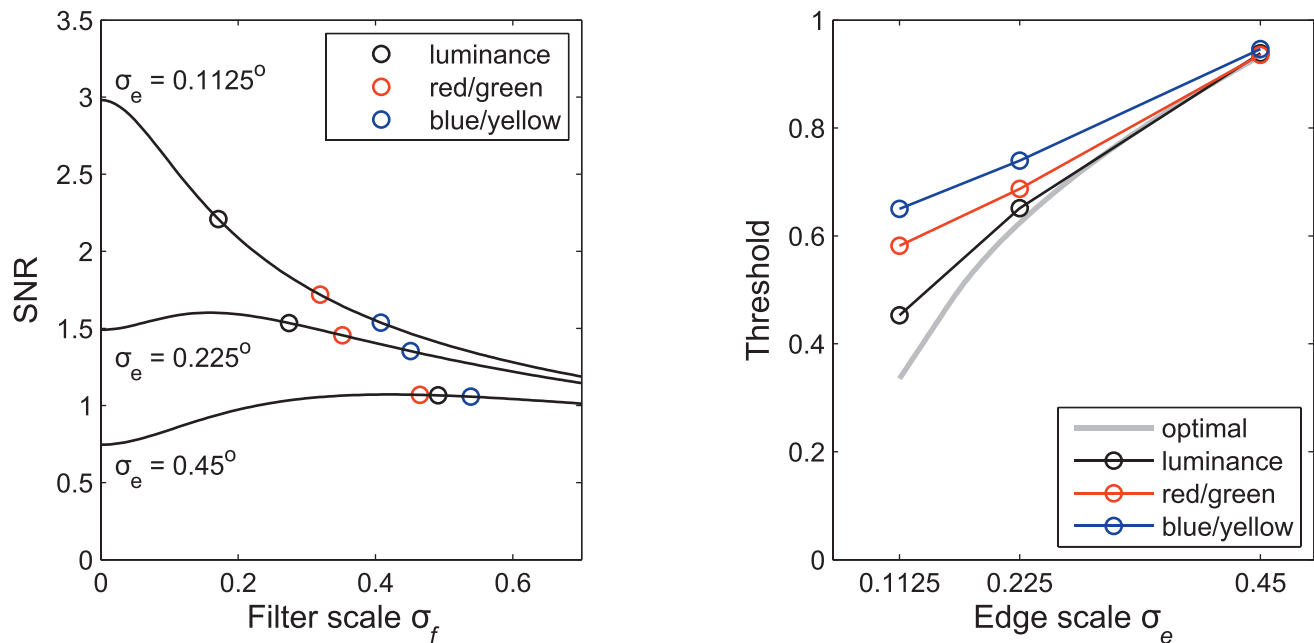


Figure 10. The left panel shows the theoretical signal-to-noise ratio (SNR) for derivative of Gaussian filters applied to the task of detection Gaussian blurred edges. The lines give the signal to noise ratio as a function of  $\sigma_f$  for the three different edge scales used in this study. The circles plot the SNR for the best-fit derivative of Gaussian for the average classification images from Figure 7. These SNR values can be inverted to give theoretical detection thresholds, shown in the right panel. The relationship between edge scale  $\sigma_e$  and threshold is approximately linear for the filter scales found in Figure 7. The gray curve shows the theoretical threshold for an optimal edge detector.

of the V1 receptive fields that detect the stimuli, and these would be the ones most sensitive to the stimulus. For red/green color, the peak spatial frequencies are 0.78, 0.71, and 0.53  $c/^\circ$  for the same three edge widths.

## Conclusions

We have used edge stimuli with the same spatial parameters but different cone contrasts to look for edge detectors in human vision. The existence of luminance edge detectors is not in dispute, so the important result here is that we have obtained evidence for edge detectors that respond to isoluminant chromatic edges. These chromatic edge detectors appear broadly similar to luminance edge detectors. However, the width of the chromatic edge detectors is larger than the luminance edge detectors. This is consistent with the lower spatial resolution of the chromatic system. This excess width of the chromatic edge detectors means that they are less effective at localizing chromatic edges, which is consistent with “capture” of chromatic edges by nearby luminance edges.

For three of our observers, the color classification images are consistent with an edge detector which is applied to the cone contrast signal, just like the luminance system. However, one of our observers does not seem to detect chromatic edges this way, and instead

only responds reliably when there is a hue change at the edge. The cause of this is not clear from our data.

The cone inputs to the chromatic edge detectors have not been determined by these experiments. They could be pure chromatic edge detectors, with no luminance sensitivity, or they could be mixed color-luminance edge detectors, as found in primate V1 (Johnson, Hawken, & Shapley, 2008). However, if these classification images do indeed represent neural receptive fields, they cannot be the only chromatically sensitive neurons in visual cortex. In particular, the edge detectors we have found will only signal a change in color, and not the actual colors on either side of the chromatic edge. To see these colors would require additional low pass chromatic channels.

*Keywords: color vision, edges, spatial vision, classification images*

## Acknowledgments

This project was funded by a Royal Society Travel Grant IE130877 to WHM and KTM, and funded in part by Canadian Institutes of Health Research (CIHR) grant MOP-10819 to KTM. We acknowledge the assistance of Nicole Telidis in data collection at McGill University.



Commercial relationships: none.

Corresponding author: William McIlhagga.

Email: w.h.mcilhagga@bradford.ac.uk.

Address: Bradford School of Optometry & Vision Science, University of Bradford, Bradford, UK.

## Footnote

<sup>1</sup> [From Appendix A] Our staircase has a probability correct of around 0.65, so the average number of bits is  $-(0.65 \log_2 0.65 + 0.35 \log_2 0.35)$  per response. The most informative responses are those where the increment contrast is so low that the probability correct is nearer 0.5, or when the observer makes a genuine mistake when the probability correct is high.

## References

- Ahumada, A. J. (1996). Perceptual classification images from Vernier acuity masked by noise. *Perception*, 25(1\_suppl), 2–2, <https://doi.org/10.1068/v96l0501>.
- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6), 716–723, <https://doi.org/10.1109/TAC.1974.1100705>.
- Anstis, S. M., & Cavanagh, P. (1983). A minimum motion technique for judging equiluminance. In L. T. Sharpe & J. D. Mollon (Eds.), *Colour vision: Psychophysics and physiology* (pp. 155–166). London: Academic Press.
- Barrow, H. G., & Tenenbaum, J. M. (1981). Interpreting line drawings as three-dimensional surfaces. *Artificial Intelligence*, 17(1), 75–116, [https://doi.org/10.1016/0004-3702\(81\)90021-7](https://doi.org/10.1016/0004-3702(81)90021-7).
- Beard, B. L., & Ahumada, A. J. (1998). Technique to extract relevant image features for visual tasks. In *Proceedings of SPIE* (pp. 79–85). San Jose, CA, USA, <https://doi.org/10.1117/12.320099>.
- Beaudot, W. H. A., & Mullen, K. T. (2005). Orientation selectivity in luminance and color vision assessed using 2-d band-pass filtered spatial noise. *Vision Research*, 45(6), 687–696, <https://doi.org/10.1016/j.visres.2004.09.023>.
- Billock, V. A. (2000). Neural acclimation to 1/f spatial frequency spectra in natural images transduced by the human visual system. *Physica D: Nonlinear Phenomena*, 137(3), 379–391, [https://doi.org/10.1016/S0167-2789\(99\)00197-9](https://doi.org/10.1016/S0167-2789(99)00197-9).
- Bradley, A., Switkes, E., & De Valois, K. (1988). Orientation and spatial frequency selectivity of adaptation to color and luminance gratings. *Vision Research*, 28(7), 841–856, [https://doi.org/10.1016/0042-6989\(88\)90031-4](https://doi.org/10.1016/0042-6989(88)90031-4).
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, 10(4), 433–436, <https://doi.org/10.1163/156856897X00357>.
- Burnham, K. P., & Anderson, D. R. (2004). Multi-model inference. *Sociological Methods & Research*, 33(2), 261–304, <https://doi.org/10.1177/0049124104268644>.
- Burton, G. J., & Moorhead, I. R. (1987). Color and spatial structure in natural scenes. *Applied Optics*, 26(1), 157–170, <https://doi.org/10.1364/AO.26.000157>.
- Canny, J. (1986). A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(6), 679–698, <https://doi.org/10.1109/TPAMI.1986.4767851>.
- Cole, G. R., & Hine, T. (1992). Computation of cone contrasts for color vision research. *Behavior Research Methods, Instruments, & Computers*, 24(1), 22–27, <https://doi.org/10.3758/BF03203465>.
- Cole, G. R., Hine, T., & McIlhagga, W. (1993). Detection mechanisms in L-, M-, and S-cone contrast space. *Journal of the Optical Society of America A*, 10(1), 38, <https://doi.org/10.1364/JOSAA.10.000038>.
- Donoho, D. L., Johnstone, I. M., Kerkycharian, G., & Picard, D. (1995). Wavelet shrinkage: Asymptopia. *Journal of the Royal Statistical Society, Series B*, 371–394.
- Field, D. J. (1987). Relations between the statistics of natural images and the response properties of cortical cells. *Journal of the Optical Society of America. A, Optics and Image Science*, 4(12), 2379–2394, <https://doi.org/10.1364/JOSAA.4.002379>.
- Friedman, J., Hastie, T., & Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *Journal of Statistical Software*, 33(1), 1–22, <https://doi.org/10.18637/jss.v033.i01>.
- Georgeson, M. A., May, K. A., Freeman, T. C. A., & Hesse, G. S. (2007). From filters to features: Scale-space analysis of edge and blur coding in human vision. *Journal of Vision*, 7(13):7, 1–21, <https://doi.org/10.1167/7.13.7>. [PubMed] [Article]
- Gheiratmand, M., Meese, T. S., & Mullen, K. T. (2013). Blobs versus bars: Psychophysical evidence supports two types of orientation response in human color vision. *Journal of Vision*, 13(1):2, 1–13, <https://doi.org/10.1167/13.1.2>. [PubMed] [Article]
- Gheiratmand, M., & Mullen, K. T. (2014). Orientation tuning in human colour vision at detection thresh-

- old. *Scientific Reports*, 4, <https://doi.org/10.1038/srep04285>.
- Hansen, T., & Gegenfurtner, K. R. (2009). Independence of color and luminance edges in natural scenes. *Visual Neuroscience*, 26(1), 35–49, <https://doi.org/10.1017/S0952523808080796>.
- Hosmer, D. W., & Lemeshow, S. (1980). Goodness of fit tests for the multiple logistic regression model. *Communications in Statistics - Theory and Methods*, 9(10), 1043–1069, <https://doi.org/10.1080/03610928008827941>.
- Humanski, R. A., & Wilson, H. R. (1992). Spatial frequency mechanisms with short-wavelength-sensitive cone inputs. *Vision Research*, 32(3), 549–560, [https://doi.org/10.1016/0042-6989\(92\)90247-G](https://doi.org/10.1016/0042-6989(92)90247-G).
- Humanski, R. A., & Wilson, H. R. (1993). Spatial-frequency adaptation: Evidence for a multiple-channel model of short-wavelength-sensitive-cone spatial vision. *Vision Research*, 33(5–6), 665–675, [https://doi.org/10.1016/0042-6989\(93\)90187-2](https://doi.org/10.1016/0042-6989(93)90187-2).
- Ihaka, R., & Gentleman, R. (1996). R: A language for data analysis and graphics. *Journal of Computational and Graphical Statistics*, 5(3), 299–314, <https://doi.org/10.1080/10618600.1996.10474713>.
- Johnson, E. N., Hawken, M. J., & Shapley, R. (2008). The orientation selectivity of color-responsive neurons in macaque V1. *Journal of Neuroscience*, 28(32), 8096–8106, <https://doi.org/10.1523/JNEUROSCI.1404-08.2008>.
- Johnson, E. N., & Mullen, K. T. (2016). Color in the cortex. In *Human color vision* (pp. 189–217). Cham, Switzerland: Springer, [https://doi.org/10.1007/978-3-319-44978-4\\_7](https://doi.org/10.1007/978-3-319-44978-4_7).
- Kim, Y. J., Reynaud, A., Hess, R. F., & Mullen, K. T. (2017). A normative data set for the clinical assessment of achromatic and chromatic contrast sensitivity using a qCSF approach. *Investigative Ophthalmology & Visual Science*, 58(9), 3628–3636, <https://doi.org/10.1167/iovs.17-21645>.
- Kingdom, F. A. A. (2003). Color brings relief to human vision. *Nature Neuroscience*, 6(6), 641–644, <https://doi.org/10.1038/nn1060>.
- Kleiner, M., Brainard, D., & Pelli, D. (2007). What's new in Psychtoolbox-3? In *Perception ECVF Abstract Supplement (Vol. 36)*.
- Knoblauch, K., & Maloney, L. T. (2008). Estimating classification images with generalized linear and additive models. *Journal of Vision*, 8(16):10, 1–19, <https://doi.org/10.1167/8.16.10>. [PubMed] [Article]
- Krauskopf, J., Williams, D. R., & Heeley, D. W. (1982). Cardinal directions of color space. *Vision Research*, 22(9), 1123–1131, [https://doi.org/10.1016/0042-6989\(82\)90077-3](https://doi.org/10.1016/0042-6989(82)90077-3).
- Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. In F. Pereira, C. J. C. Burges, L. Bottou, & K. Q. Weinberger (Eds.), *Advances in neural information processing systems 25* (pp. 1097–1105). Red Hook, NY: Curran Associates, <https://doi.org/10.1145/3065386>.
- Kuss, O. (2002). Global goodness of fit tests in logistic regression with sparse data. *Statistics in Medicine*, 21(24), 3789–3801, <https://doi.org/10.1002/sim.1421>.
- Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *The Journal of the Acoustical Society of America*, 49(2B), 467–477, <https://doi.org/10.1121/1.1912375>.
- Lindeberg, T. (1998). Feature detection with automatic scale selection. *International Journal of Computer Vision*, 30(2), 79–116, <https://doi.org/10.1023/A:1008045108935>.
- Livingstone, M. S., & Hubel, D. H. (1984). Anatomy and physiology of a color system in the primate visual cortex. *Journal of Neuroscience*, 4(1), 309–356.
- Livingstone, M. S., & Hubel, D. H. (1987). Psychophysical evidence for separate channels for the perception of form, color, movement, and depth. *Journal of Neuroscience*, 7(11), 3416–3468.
- Losada, M. A., & Mullen, K. T. (1994). The spatial tuning of chromatic mechanisms identified by simultaneous masking. *Vision Research*, 34(3), 331–341, [https://doi.org/10.1016/0042-6989\(94\)90091-4](https://doi.org/10.1016/0042-6989(94)90091-4).
- Losada, M. A., & Mullen, K. T. (1995). Color and luminance spatial tuning estimated by noise masking in the absence of off-frequency looking. *Journal of the Optical Society of America A*, 12(2), 250–260, <https://doi.org/10.1364/JOSAA.12.000250>.
- Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. Cambridge, MA: MIT Press.
- McCullagh, P., & Nelder, J. A. (1989). *Generalized linear models* (Vol. 37). Boca Raton, FL: CRC Press.
- McIlhagga, W. (2011). The canny edge detector revisited. *International Journal of Computer Vision*, 91, 251–261. <http://dx.doi.org/10.1007/s11263-010-0392-0>.
- McIlhagga, W. (2016). penalized: A MATLAB toolbox for fitting generalized linear models with penalties. *Journal of Statistical Software, Articles*, 72(6), 1–21, <https://doi.org/10.18637/jss.v072.i06>.
- McIlhagga, W., & May, K. A. (2012). Optimal edge filters explain human blur detection. *Journal of Vision*, 12(10):9, 1–13, <https://doi.org/10.1167/12.10.9>. [PubMed] [Article]

- Mineault, P. J., Barthelmé, S., & Pack, C. C. (2009). Improved classification images with sparse priors in a smooth basis. *Journal of Vision*, 9(10):17, 1–24, <https://doi.org/10.1167/9.10.17>. [PubMed] [Article]
- Mullen, K. T. (1985). The contrast sensitivity of human colour vision to red-green and blue-yellow chromatic gratings. *The Journal of Physiology*, 359(1), 381–400, <https://doi.org/10.1113/jphysiol.1985.sp015591>.
- Mullen, K. T., & Losada, M. A. (1999). The spatial tuning of color and luminance peripheral vision measured with notch filtered noise masking. *Vision Research*, 39(4), 721–731, [https://doi.org/10.1016/S0042-6989\(98\)00171-0](https://doi.org/10.1016/S0042-6989(98)00171-0).
- Murray, R. F. (2011). Classification images: A review. *Journal of Vision*, 11(5):2, 1–25, <https://doi.org/10.1167/11.5.2>. [PubMed] [Article]
- Murray, R. F., Bennett, P. J., & Sekuler, A. B. (2002). Optimal methods for calculating classification images: Weighted sums. *Journal of Vision*, 2(1):6, 79–104, <https://doi.org/10.1167/2.1.6>. [PubMed] [Article]
- Olmos, A., & Kingdom, F. A. A. (2004a). A biologically inspired algorithm for the recovery of shading and reflectance images. *Perception*, 33(12), 1463–1473, <https://doi.org/10.1068/p5321>.
- Olmos, A., & Kingdom, F. A. A. (2004b). McGill Calibrated Colour Image Database. Retrieved from <http://tabby.vision.mcgill.ca/html/welcome.html>
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, 10(4), 437–442, <https://doi.org/10.1163/156856897X00366>.
- Reisbeck, T. E., & Gegenfurtner, K. R. (1998). Effects of contrast and temporal frequency on orientation discrimination for luminance and isoluminant stimuli. *Vision Research*, 38(8), 1105–1117, [https://doi.org/10.1016/S0042-6989\(97\)00240-X](https://doi.org/10.1016/S0042-6989(97)00240-X).
- Sankeralli, M. J., & Mullen, K. T. (1996). Estimation of the L-, M-, and S-cone weights of the postreceptoral detection mechanisms. *Journal of the Optical Society of America A*, 13(5), 906–915, <https://doi.org/10.1364/JOSAA.13.000906>.
- Schluppeck, D., & Engel, S. A. (2002). Color opponent neurons in V1: A review and model reconciling results from imaging and single-unit recording. *Journal of Vision*, 2(6):5, 480–492, <https://doi.org/10.1167/2.6.5>. [PubMed] [Article]
- Shapley, R. M., & Hawken, M. J. (2011). Color in the cortex: Single- and double-opponent cells. *Vision Research*, 51(7), 701–717, <https://doi.org/10.1016/j.visres.2011.02.012>.
- Shapley, R. M., & Tolhurst, D. J. (1973). Edge detectors in human vision. *The Journal of Physiology*, 229(1), 165–183.
- Stromeyer, C. F., Cole, G. R., & Kronauer, R. E. (1985). Second-site adaptation in the red-green chromatic pathways. *Vision Research*, 25(2), 219–237, [https://doi.org/10.1016/0042-6989\(85\)90116-6](https://doi.org/10.1016/0042-6989(85)90116-6).
- Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 58(1), 267–288.
- Vimal, R. L. P. (1997). Orientation tuning of the spatial-frequency-tuned mechanisms of the Red-Green channel. *Journal of the Optical Society of America A*, 14(10), 2622–2632, <https://doi.org/10.1364/JOSAA.14.002622>.
- Webster, M. A., Switkes, E., & Valois, K. K. D. (1990). Orientation and spatial-frequency discrimination for luminance and chromatic gratings. *Journal of the Optical Society of America A*, 7(6), 1034–1049, <https://doi.org/10.1364/JOSAA.7.001034>.
- Wuerger, S. M., & Morgan, M. J. (1999). Input of long- and middle-wavelength-sensitive cones to orientation discrimination. *Journal of the Optical Society of America A*, 16(3), 436–442, <https://doi.org/10.1364/JOSAA.16.000436>.

## Appendix A

From Equation 2, the decision variable is

$$d_i = \sum_x t_x (C_i(x) + n_i(x) - m_i(x))$$

The difference  $d_i$  affects the probability of a correct response. The probability  $p_i$  that the observer is correct on trial  $i$  is a logistic function of the decision variable  $d_i$ :

$$p_i = (1 + \exp(-d_i))^{-1}$$

The observer's actual response is recorded by a variable  $r_i$ , which is 1 if the observer is correct on trial  $i$ , and 0 if they are incorrect. The log-likelihood of the observer responses, given the template  $t_y$  is then

$$L = \sum_i r_i \log p_i + (1 - r_i) \log (1 - p_i)$$

The log-likelihood  $L$  is a function of the template weights  $t_y$ , so they can be estimated by maximum likelihood. These three equations define a logistic regression, which is a form of generalized linear model (McCullagh & Nelder, 1989).

It might be expected that with 1,800 trials it should be fairly easy to estimate the 150 parameters  $t_x$  that form the template. However, these 1,800 trials do not



contain much information - on average, only about 0.93 bits<sup>1</sup> per trial. Our response data thus contains less than 1,500 bits of information, which is the same as 150 measurements of  $d_i$  at a precision of 10 bits each. Thus our logistic regression on 1,800 data points is, from an information angle, equivalent to a linear regression with only 150 measurements of  $d_i$ . Since we are trying to estimate 150 values for the template  $t_x$ , over-fitting is pretty much guaranteed, and standard logistic regression fails to yield any sensible estimate of the template  $t_x$ .

One way to avoid over-fitting is to constrain the template so it can be encoded with fewer bits. There are many possible choices for the constraint, with no clear criterion for choosing one over the other. Here, we constrain the template so that it is sparse in the wavelet domain. This is a similar approach to that taken by (Mineault et al., 2009). One reason for choosing this constraint is that it leads to estimates of the template that are guaranteed to be at least as smooth as the true template shape, although the definition of “smooth” is slightly esoteric (Donoho, Johnstone, Kerkycharian, & Picard, 1995), and does not always yield what the eye considers smooth.

A wavelet basis is a set of mutually orthogonal vectors  $W_{y,1}, W_{y,2}, \dots$  each with unit length. The wavelet coefficients  $w_1, w_2, \dots$  of the template are given by the dot-products  $w_k = \sum_y W_{y,k} t_y$ , so that

$$t_y = \sum_k w_k W_{y,k}$$

Sparseness is enforced by adding a penalty on the wavelet coefficients to the logistic regression. The most common penalty is the  $L_1$ , or LASSO, penalty (Tibshirani, 1996),  $\lambda \sum_k |w_k|$ . Combining this with the likelihood, the estimated template is the set of values  $t_y$  which maximizes

$$L = \sum_i (r_i \log p_i + (1 - r_i) \log(1 - p_i)) - \lambda \sum_k |w_k|$$

The best penalty weight  $\lambda$  was chosen by 5-fold cross validation. There are efficient packages available for this maximization. In R (Ihaka & Gentleman, 1996), the standard package for this is glmnet (Friedman, Hastie, & Tibshirani, 2010). The one used here is a Matlab package described in McIlhagga (2016). The package takes about 20 s to estimate the template and do five-fold cross-validation, for four different wavelets, for one observer, using a Core i5-4590 3.3GHz processor, so it is quite a bit faster than first order methods (Mineault et al., 2009).

The only remaining question is which wavelet, from the many possible ones available, should be used? There are, unfortunately, no good criteria for choosing one wavelet basis over another. Instead, the aforementioned fitting procedure was performed for five

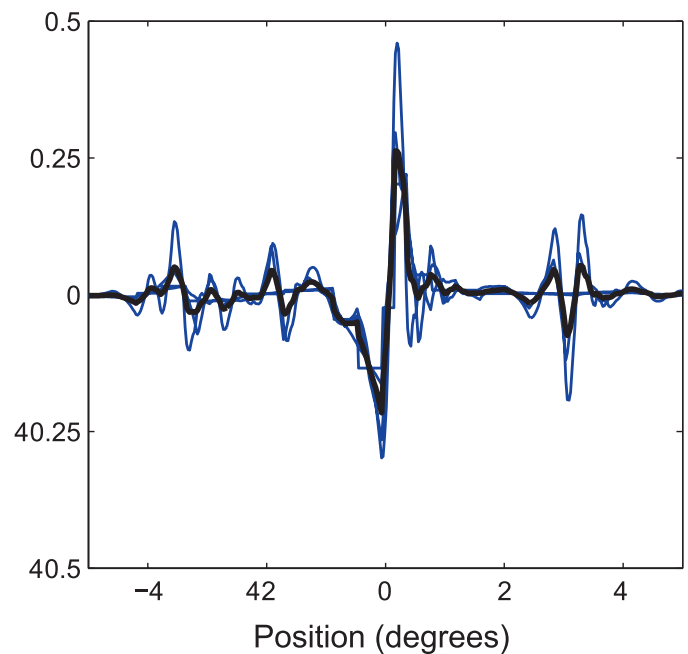


Figure 11. Multiple wavelet estimates (thin blue lines) and average (thick black line) for observer E,  $\sigma = 0.225^\circ$ .

different wavelet bases (Haar, Daubechies 2 and 3 tap wavelets, and symlet 4 and 6) and the template estimates from each wavelet were then averaged. An example of individual wavelet estimates for the various bases, together with their average, is shown in Figure 11.

This is a simple form of model averaging. Most advice on model averaging suggests using the exponent of the negative Akaike information criterion (Akaike, 1974) as the model average weight (Burnham & Anderson, 2004). However, bootstrapping shows that the AIC has a sample standard deviation of about 30 for our data, but using the exponent of the negative AIC means that differences as small as 5 can have a major impact on the weightings. For this reason, we avoided using such unreliable weightings when averaging, and calculated simple unweighted averages.

## Appendix B

Here, we derive the formulas related to the optimal (signal-to-noise) edge detection filter for a Gaussian blurred edge in Gaussian blurred noise. The localization criterion has been ignored, as it may not have much effect on the filter shape, and is, in any case, difficult to manage.

We begin with a Gaussian blurred step edge, with blurring filter size  $\sigma_e$ , in Gaussian blurred brown noise, with blurring filter size  $\sigma_n$ . The first step is to take the derivative, to yield a Gaussian blob, with width  $\sigma_e$  in



Gaussian blurred white noise, with blurring filter width  $\sigma_n$ .

The optimal filter is easiest to work out in the Fourier domain. The Fourier transform of the Gaussian blob is

$$\exp(-2\pi^2\sigma_e^2f^2)$$

And the power spectrum of the blurred noise is

$$(\exp(-2\pi^2\sigma_n^2f^2))^2 = \exp(-4\pi^2\sigma_n^2f^2)$$

In the Fourier domain, the optimal filter with respect to signal-to-noise is given by the Fourier transform of the Gaussian blob divided by the power spectrum of the noise, namely

$$\exp(-2\pi^2(\sigma_e^2 - 2\sigma_n^2)f^2)$$

The optimal detector is thus the derivative of this. If instead we use a Gaussian filter with width  $\sigma_f$ , we can work out the signal-to-noise ratio as follows. The signal

S is given by

$$\begin{aligned} S &= \int \exp(-2\pi^2\sigma_e^2f^2) \exp(-2\pi^2\sigma_f^2f^2) df \\ &= \frac{1}{\sqrt{2\pi}\sqrt{\sigma_e^2 + \sigma_f^2}} \end{aligned}$$

The noise variance  $N^2$  is given by

$$\begin{aligned} N^2 &= \int \exp(-4\pi^2\sigma_n^2f^2) \exp(-2\pi^2\sigma_f^2f^2)^2 df \\ &= \frac{1}{\sqrt{2\pi}\sqrt{\sigma_n^2 + \sigma_f^2}} \end{aligned}$$

From this, the signal-to-noise ratio  $S/N$  is proportional to

$$S/N \propto \sqrt{\frac{\sqrt{\sigma_n^2 + \sigma_f^2}}{\sigma_e^2 + \sigma_f^2}}$$